

Japanese Patent Laid-open Publication No. 2002-344957 A

Publication date : November 29, 2002

Applicant : HITACHI KOKUSAI ELECTRIC INC

Title : IMAGE MONITORING SYSTEM

5

(57) [Abstract]

[Object] To provide an image monitoring system that can automatically move a video camera toward a particular surveillance target.

[Means] A plurality of microphones 1 to 5 placed with different azimuths  
10 within a horizontal plane as a microphone group 102 are used to identify the azimuth of a sound source (intruding object) S, and a electric pan head 103 is controlled to automatically direct a video camera 101 to the direction of the sound source S.

[Effect] When the object emitting sound enters a surveillance target region, the  
15 video camera is automatically pointed to the direction of the object, and the image and sound of the intruding object can reliably be obtained without the need for checking a monitor at all times.

[0045] The transmission of a sound signal of this embodiment will be  
20 explained next. In the image monitoring system of the present invention, audio communication is necessary in some cases between a monitoring site and a monitoring center and at that time, two-way communication, such as a general telephone, is preferable,. However, in order to make this possible, a transmission path is required. If it is possible to commonly use a dedicated  
25 transmission path existing between the monitoring site and the monitoring

**THIS PAGE BLANK (USPTO)**

center, which is used for transmitting image data or control signal for monitoring, it is more advantageous in terms of costs.

**THIS PAGE BLANK (USPTO)**

(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2002-344957  
(P2002-344957A)

(43)公開日 平成14年11月29日 (2002.11.29)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テ-マコ-ト <sup>*</sup> (参考)
H 0 4 N 7/18		H 0 4 N 7/18	G 5 C 0 2 2
G 0 8 B 13/196		G 0 8 B 13/196	E 5 C 0 5 4
25/00	5 1 0	25/00	5 C 0 8 4
G 1 0 L 15/00		H 0 4 N 5/222	5 1 0 M 5 C 0 8 7
			B 5 D 0 1 5
		審査請求 未請求 請求項の数 2 O L (全 9 頁)	最終頁に統く

(21)出願番号 特願2001-141881(P2001-141881)

(22)出願日 平成13年5月11日(2001.5.11)

(71)出願人 000001122

株式会社日立国際電気

東京都中野区東中野三丁目14番20号

(72)発明者 住吉 正紀

東京都小平市御幸町32番地 株式会社日立  
国際電気小金井工場内

(74)代理人 100078134

弁理士 武 顕次郎

最終頁に統く

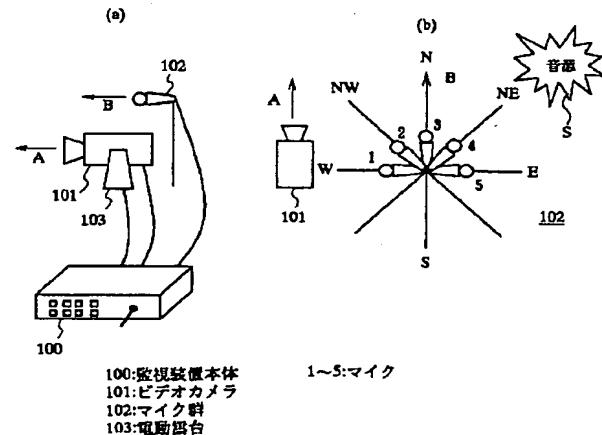
## (54)【発明の名称】 画像監視装置

## (57)【要約】

【課題】 特定の監視対象に向けて自動的にビデオカメラを動かすことができるようとした画像監視装置を提供すること。

【解決手段】 マイク群102として、音源(侵入物体)Sの方位を、水平面内で異なる方位に向けて配置した複数のマイク1～5を用いて特定し、電動雲台103を制御して、ビデオカメラ101を音源Sの方向に自動的に向けるようにしたもの。

【効果】 音声を発している物体が監視対象領域に入れば、それに向けてビデオカメラが自動的に動かされるので、常時、モニタの監視を要することなく、確実に侵入物体の画像と音声を得ることができる。



よる撮像視野を変え、必要な場所が任意に撮像できるようになっていた。

【0006】ここで、ビデオカメラを動かして撮像視野を変更する方法の場合、オペレータが操作器などを用い、手動でビデオカメラを操作するのが一般的であるが、しかし、この方法では、オペレータが當時、或いは必要なとき、モニタを監視している必要がある。

【0007】そこで、ビデオカメラを自動的に操作する方法として、例えばビデオカメラの撮像方向を、予め何点かプリセット方向として設定しておき、これらのプリセット位置を順次、自動的に移し替えてゆく方法、或いは侵入物体をビデオカメラで撮像した画像データから画像認識し、認識した物体が追尾されるようにビデオカメラを動かす方法などが、従来から提案されている。

#### 【0008】

【発明が解決しようとする課題】上記従来技術は、監視対象の特定に充分な配慮がされているとは言えず、ビデオカメラによる対象画像の撮像に問題があった。すなわち、まず、プリセット位置の移し替えによる方法では、何らかの事象が発生した時点で、ビデオカメラがその方向に向いていない場合があり、必要な画像のモニタに問題が生じてしまう。

【0009】また、画像認識によりビデオカメラを追尾させる方法では、撮像対象が複数現われ、それらが異なる方向に移動した場合には、撮像対象を何れか1種に絞らざるを得ず、しかもビデオカメラの撮像視野外で発生した事象には対応できないという問題が生じてしまう。

【0010】本発明の目的は、特定の監視対象に向けて自動的にビデオカメラを動かすことができるようした画像監視装置を提供することにある。

#### 【0011】

【課題を解決するための手段】上記目的は、ビデオカメラで撮像した画像により、対象領域に侵入した物体を監視するようにした画像監視装置において、水平面内で異なる方位に向けて配置した複数のマイクロホンにより検出される音データの解析により、音源の方位を特定する手段を設け、前記ビデオカメラの撮像方向が、前記手段により特定した方位に自動的に制御されるようにして達成される。このとき、前記マイクロホンで検出された音データを音声認識し、文字コードデータに変換する手段が設けられているようにしても良い。

#### 【0012】

【発明の実施の形態】以下、本発明による画像監視装置について、図示の実施の形態により詳細に説明する。図1は、本発明の一実施形態が適用されたシステムの一例で、まず、図の(a)に示すように、このシステムの全体は、監視装置本体100とビデオカメラ101、マイク群102、それに電動雲台103で構成されている。

【0013】まず、監視装置本体100について、図2 50 により説明する。

#### 【特許請求の範囲】

【請求項1】ビデオカメラで撮像した画像により、対象領域に侵入した物体を監視するようにした画像監視装置において、

水平面内で異なる方位に向けて配置した複数のマイクロホンにより検出される音データの解析により、音源の方位を特定する手段を設け、

前記ビデオカメラの撮像方向が、前記手段により特定した方位に自動的に制御するように構成したことを特徴とする画像監視装置。

【請求項2】請求項1に記載の発明において、前記マイクロホンで検出された音データを音声認識し、文字コードデータに変換する手段が設けられていることを特徴とする画像監視装置。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】本発明は、画像による侵入物体の監視や定点観測に使用される画像監視装置に係り、特に監視範囲や観測範囲が広い場合に好適な画像監視装置に関する。

##### 【0002】

【従来の技術】近年、侵入物体の監視や定点観測などに、ビデオカメラで撮像した画像を用いた監視システムが広く採用されるようになっている。また、近年は、モバイル画像表示機器やインターネットなどの普及に伴って、希望する領域の画像を、遠隔地の監視センタで監視したり観測できるようにしたシステムも提供されるようになっている。

【0003】このとき、従来の画像監視装置では、ビデオカメラの画像信号や音声信号をそのまま監視センタに伝送する方式と、符号化(圧縮)して伝送する方式があり、ここで、さらに、これらの信号の一部若しくは全部を記録する場合がある。そして、監視センタなどの伝送先では、そのまま、或いは復号化(符号化伝送方式の場合)して、画像をモニタに表示し、音声をスピーカにより出力する。

【0004】ここで符号化伝送方式による場合、通常、画像の圧縮にはJ E P G方式やM P E G方式が、音声にはW A V E方式などが採用され、また、伝送路には、画像信号線や無線などが用いられている。更に、記憶媒体には、I C メモリやハードディスクが使用され、ビデオカメラの制御は、シリアル信号を用いて行うのが一般的であり、このとき、サーバ機能により録画データの配信が行われる場合もある。

【0005】そこで、このような監視装置において、監視対象が広範囲にわたる際には、従来から、複数台のカメラを用いて異なった場所が撮影できるようにしたり、ビデオカメラを遠隔操作可能な雲台に取付け、カメラのパン(水平旋回)操作、チルト(仰角変更)操作、ズーミング(レンズの画角変更)操作などにより、ビデオカメラに

【0014】この図2において、CPU200は、バス201を介して、これに接続されている各々の機能を制御し、画像監視に必要な各種の処理を実行する。次に、画像圧縮／伸張部202は、ビデオカメラ102から入力される画像信号の圧縮処理と、圧縮された画像データの伸張処理を実行する。画像認識部203には、ビデオカメラ101から画像信号が入力され、侵入物体の検出に必要な処理を、予め設定してある画像処理手法により実行する。

【0015】音声合成部204では、文字コードデータを入力して音声合成し、音声信号に変換する処理を実行する。音声圧縮／伸張部205は、マイク群102から入力される音声信号の圧縮処理と、圧縮された音声データの伸張処理を実行する。音声認識部206では、入力された音声データを処理し、文字データに変換する処理を実行する。メモリ207は、監視装置本体100のプログラムを実行可能な状態で格納する働きをする。

【0016】カメラ制御部208は、ビデオカメラ102及び電動雲台103と制御信号の授受を行ない、ビデオカメラ102の方向制御、フォーカス制御、それにズーム制御などに必要な処理を実行する。ネットワーク制御部209は、それが接続されているインターネットやLANなどのネットワークNのプロトコルを認識し、ネットワークNとデータの送受信に必要な処理を実行する。

【0017】記録装置210は、画像データと音声データ、文字コードデータ、更には監視装置本体100で扱う各種の情報を記録するもので、記録されたデータはCPU200の要求により選択的に読み出される。このため、ハードディスクやDVD、メモリなどランダムアクセスが可能な媒体であれば何であってもよい。入力装置211は、例えばキーボードなど、オペレータとのインターフェース用で、例えば後述する動作モードの切替えなど各種のデータの設定や指示の入力に使用される。

【0018】図1に戻り、ビデオカメラ101は、所定の画角のレンズ、或いはズームレンズが装備しており、最大写角の状態で、監視対象とする領域の上下方向では必要な範囲の全体を含み、水平方向では一部だけの撮像が可能なCCDカメラなどで構成され、必要に応じてカラーカメラとモノクロカメラが選択される。

【0019】ここで、図示してある矢印Aは、ビデオカメラ101の向き(撮像中心方向)が監視対象とする領域の左右方向の中心を向いているときの方向を表わす。

【0020】次に、マイク群102は、ビデオカメラ101の近傍に、水平面内で異なる方位に向けて配置した複数個(ここでは5個)のマイク1～5で構成されている。なお、ここで“マイク”とは、マイクロホンの略称である。そして、これらのマイク1～5は、図1の(b)から明らかのように、ビデオカメラ1の向きAと同じ方向Bを中心として、水平面内の左右に略等角度で均等に

配置してある。

【0021】次に、電動雲台103は、ビデオカメラ102を保持し、このとき、少なくとも方位角については遠隔制御が可能なカメラ架台で構成され、カメラ制御部208(図2)から供給されるカメラ制御信号に応じてビデオカメラ101をパン移動させ、方位角を遠隔制御する働きをする。

【0022】次に、この実施形態の動作について説明する。図3は、この実施形態の動作を説明するためのフローチャートで、この処理はCPU200により実行され、この処理が開始(Start)されると、まず、映像信号と音声信号をビデオカメラ101とマイク群102から入力し(処理300)、画像圧縮／伸張部202及び音声圧縮／伸張部205で圧縮処理する(処理301)。

【0023】そして、この後は、動作モードに応じて処理が分岐される(処理302)。このときの動作モードは、入力装置211(図2)からCPU200に入力される動作モード指令、又は監視センタからCPU200に供給される動作モード指令によって設定されるもので、この実施形態では、モニタモードMAと検知モードMB、それにカメラ追従モードの3種の動作モードがある。

【0024】まず、ここで、いま、モニタモードMAに設定されていたとする。このモニタモードMAは、通常の監視動作用の動作モードで、このため、処理301で圧縮処理された映像データと音声データは、この後、記録装置210に記録格納されると共に、ネットワーク制御部209を介してネットワークN上に配信される(処理303)。

【0025】この結果、ネットワークNに接続されている監視センタ(図示されていない)では、監視対象領域からビデオカメラ101により撮像された画像と、同じく監視対象領域からマイク群102により集音された音声を、そこにある所定のモニタ用機器(画像表示装置やスピーカなど)で再生することができ、画像と音声による監視が得られることになる。

【0026】次に、検知モードMBに設定されていたとする。このモードMBは、ビデオカメラ1で撮像した画像データから画像認識により侵入物体を検出する動作モードで、このときは、画像認識部203に画像データが供給され、ここで、予め設定してあるアルゴリズムに従って侵入物体の検知が行なわれ(処理304)、この後、該当する画像データと音声データが選択的に記録格納されると共に、ネットワークN上に配信される(処理303)。

【0027】なお、画像データと音声データについては、データ圧縮を行わない状態(非圧縮状態)で記録格納され、配信を行うようにしてもよい。また、このとき、マイク群102には複数個のマイク1～5があるので、これらの音声は、夫々別々に記録格納され、配信される

ようにしてもよい。

【0028】次に、カメラ追従モードMCに設定されたとする。このモードMCは、図1(b)に、音源Sとして示したように、ビデオカメラ101による監視対象領域の中に、例えば自動車など、音を発生する侵入物体が現われたとき、この侵入物体の方向が自動的に検出され、この方向にビデオカメラ101の撮像方向が自動的に追尾される動作を与えるモードのことである。

【0029】このカメラ追従モードMCときは、まず、各マイク1～5から入力された音声データをCPU200で解析する(処理305)。次に、この解析結果に基づいて音源の方向を検出し特定する(処理306)。そして、この特定結果を、カメラ制御部208を介して電動雲台102に供給する(処理307)。

【0030】この結果、ビデオカメラ101の撮像方向は、音源S、つまり侵入物体が存在する方向に自動的に向けられ、従って、ビデオカメラ101からは、侵入物体が被写体となっている画像の信号が常に確実に得られることになり、容易に侵入物体が監視できることになる。

【0031】次に、このときのCPU200による処理、すなわち、処理305と処理306による音声データの解析処理と音源方向の特定処理について、図4により具体的に説明する。この図4は、各マイク1～5で検出される音声信号の波形を並べて示したもので、図において、各表示区画401～405が、各マイク1～5の音声信号の表示区画に夫々対応しており、夫々の区画401～405に表示されている波形410～450が各マイク1～5の音声信号の波形に対応している。

【0032】ここで、各マイク1～5の水平面内の配置についてみると、既に図1(b)により示したように、ビデオカメラ1が監視対象とする領域の左右方向の中心方向Aと同じ方向Bを中心として、水平面内の左右に略等角度で均等に配置されている。

【0033】そこで、いま、この図1(b)において、例えれば方向Bを方位N(北)とすると、図示のように、マイク1は中心から方位W(西)側に位置し、以下、マイク2は方位NW(北西)側、マイク3は方位N(北)側、マイク4は方位NE(北東)側、そしてマイク5は方位E(東)側に、夫々位置していることになり、ここで、方位Nがビデオカメラ1の監視対象領域の左右方向の中心方向になる。

【0034】そうすると、監視対象とする領域内から音が発生された場合、その音源から各マイク1～5までの距離は、各マイク1～5毎に異なったものとなるようになる。そこで、この実施形態では、マイク1～5により音声信号が検出されたとき、各マイク1～5の音声信号間に現われる立上り時間の違いと、信号レベルの違いにより、音源の方位を検出し特定するようになってい る。

【0035】いま、ここで、音源Sが、図1(b)に示す位置に現われたとすると、これによる音(音声)がマイク群102により検出され、各マイク1～5から音声信号がCPU200に取り込まれる。そして、このときの音声信号が図4の波形410～450に示すようになっていたとすると、このときは、音源Sに一番近い位置にあるのがマイク4であるから、マイク4による音声信号440が最初に検出される。

【0036】そこで、この音声信号440が検出された時刻をt<sub>0</sub>とし、この後に検出される音声信号が何れのマイクによるものかを調べることにより、音源Sの方位が特定できる。この例では、図1(b)に示すように、音源Sは、マイク4が位置している方位NEにあるので、マイク3とマイク5は、音源Sから等距離にある。

【0037】そこで、マイク4の音声信号440が検出された時刻t<sub>0</sub>以降、所定の遅れ時間Δt経過後の所定の時刻t<sub>1</sub>で同時にマイク3の音声信号430とマイク5の音声信号450が検出されることになり、従って、音源Sの方向は、マイク4の方位と同じ方位NEであると特定できることになる。

【0038】以下、同様に、音源Sがマイク3の方位Nにあるときは、最初にマイク3による音声信号430が検出された後、マイク2の音声信号420とマイク4による音声信号440が同時に検出され、マイク2の方位NWにあるときは、最初にマイク2による音声信号420が検出された後、マイク1の音声信号410とマイク3による音声信号430が同時に検出されることになり、いずれの場合も音源Sの方位が特定できることになる。

【0039】次に、音源Sがマイク3の方位Nとマイク4の方位NEの間の方位、つまり方位NNE(北北東)にあるときは、音源Sからの距離がマイク3とマイク4で同じになるので、マイク3の音声信号430とマイク4の音声信号440が同時に最初に検出されることになり、音源Sは方位NNEであると特定できる。

【0040】以下、同様に、マイク1による音声信号410とマイク2による音声信号420が最初に同時に検出されたときは、音源Sは方位Wと方位NWの間、つまり方位NNW(北北西)にあり、マイク1による音声信号410とマイク2による音声信号420が最初に同時に検出されたときは、音源Sは方位Wと方位NWの間、つまり方位WNW(西北西)であると夫々特定できる。

【0041】同じく、マイク3による音声信号430とマイク4による音声信号440が最初に同時に検出されたときは、音源Sは方位Nと方位NEの間、つまり方位NNE(北北東)にあり、マイク4による音声信号440とマイク5による音声信号450が最初に同時に検出されたときは、音源Sは方位NEと方位Eの間、つまり方位ENE(東北東)であると夫々特定できる。

【0042】更に、図1(b)から明らかなように、音源

Sの方位と、音源Sから各マイク1～5までの距離は一義的に決まる。従って、この実施形態によれば、各マイク1～5で検出された音声信号間での時間差を検出することにより、上記に例示した方位以外の任意の方位にある音源Sでも、その方位を容易に、しかも充分に高い精度で特定することができる。

【0043】こうして、音源Sの方位ができたら、CPU200はカメラ制御器208を介して電動雲台103を制御し、ビデオカメラ101が音源Sの方向に向くように制御することができ、従って、侵入物体が被写体として含まれている画像をビデオカメラ101で撮影し、その画像信号を記録、配信することができ、必要な監視を確実に遂行することができる。

【0044】ここで、本発明の実施の形態としては、マイク群102により集音された音声信号のレベルが、予め設定してある一定値以上になったときだけ、音源方位によるビデオカメラ101の制御を行なうようにしてもよい。また、上記実施形態の場合、各マイク1～5による音声信号のレベルも、音源方位によって変化し、各マイク1～5の指向性によっても変化する。従って、本発明の実施形態によれば、この音声信号のレベル変化を利用して音源の方位を検出するようにすることもできる。

【0045】次に、この実施形態による音声信号の伝送について説明する。本発明が対象としている画像監視装置では、監視現場と監視センタの間で音声による通話が必要な場合があり、しかも、このとき、一般的の電話と同様、双方向通話が望ましい。しかし、このためには別途、そのための伝送路を要するが、このとき、監視現場と監視センタの間に本来の監視のための画像データや制御信号の伝送に使用されている伝送路が共用できれば、コスト的に有利である。

【0046】しかしながら、ここで、双方向通話のための音声データをそのままの形式で送信するには、かなりの伝送帯域を確保する必要があり、このため、そのまま共用した場合には、本来の伝送容量が圧迫されてしまう上、伝送量の確保に制約があった。

【0047】そこで、この実施形態では、交信する音声信号を音声認識し、これにより文字コードデータに変換し、その後、この文字コードデータを伝送し、受信先で音声合成装置により音声として取り出せるようにしたものである。このように、音声データを文字コードに変換すると、データ量がかなり少くなり、従って、音声信号の伝送に必要な伝送量の増加を最小限に抑えることができ、共用が可能になるのである。

【0048】まず、この実施形態における送信時の処理について、図5のフローチャートにより説明する。この図5に示した処理もCPU200により実行され、処理が開始されると、まず、マイク群102から音声信号を入力し(処理500)、これを音声認識部206で解析し(処理501)、次いで文字コードに変換する(処理5

02)。変換された文字コードデータは、そのときの画像データに対応させて記憶装置210に記録し、これと共に、ネットワーク制御部209を介してネットワークN上に配信される(処理503)。

【0049】次に、受信時の処理について説明する。ネットワークNからネットワーク制御部209を介して送られてきた文字コードデータ又は記憶装置210から読み出された文字コードデータは、音声合成部204に入力され、ここで音声信号に変換される。そこで、この音声合成部204から出力される音声信号をスピーカに供給することにより音声が再生される。

【0050】従って、これにより双方向での会話を行なうことができ、しかも、このとき、本来の監視用データの伝送に使用されている伝送路を用いているにも関わらず、本来の伝送容量を圧迫することなく、双方向通話に必要な伝送量が容易に確保できる。なお、この音声信号の文字コード化と音声化は、複数のマイク1～5の入力に対して行なってもよいことは言うまでもない。

#### 【0051】

【発明の効果】本発明によれば、音声を発している物体が監視対象領域に入れば、それに向けてビデオカメラが自動的に動かされるので、常時、モニタの監視を要することなく、確実に侵入物体の画像と音声を得ることができる。

#### 【図面の簡単な説明】

【図1】本発明による画像監視装置の一実施形態を示す構成図である。

【図2】本発明の一実施形態における監視装置本体のブロック構成図である。

【図3】本発明の一実施形態による監視動作処理を説明するためのフローチャートである。

【図4】本発明の一実施形態における音声信号の波形図である。

【図5】本発明の一実施形態による音声信号処理を説明するためのフローチャートである。

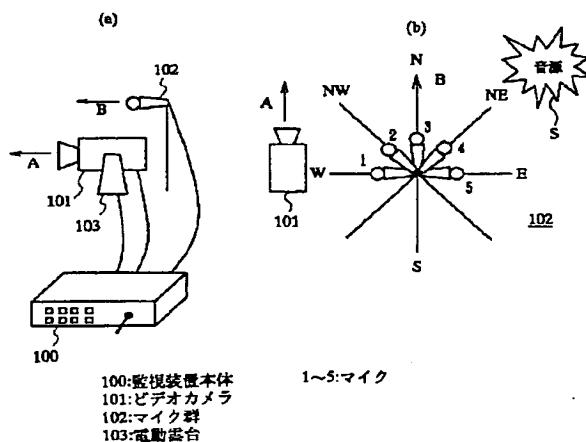
#### 【符号の説明】

1～5	マイク(マイクロホン)
100	監視装置本体
101	ビデオカメラ
102	マイク群
103	電動雲台
200	CPU
201	バス
202	画像伸張／圧縮部
203	画像認識部
204	音声合成部
205	音声圧縮／伸張部
206	音声認識部
207	メモリ
208	カメラ制御部

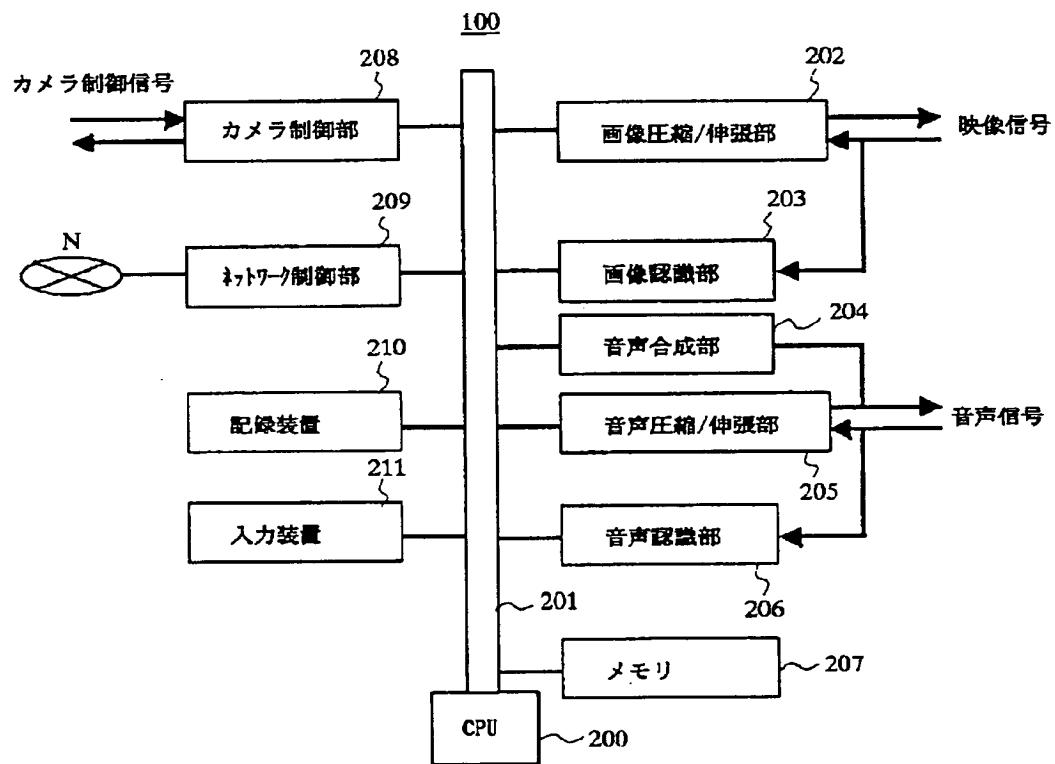
209 ネットワーク制御部  
210 記録装置  
211 入力装置

N ネットワークーク  
S 音源(侵入物体)

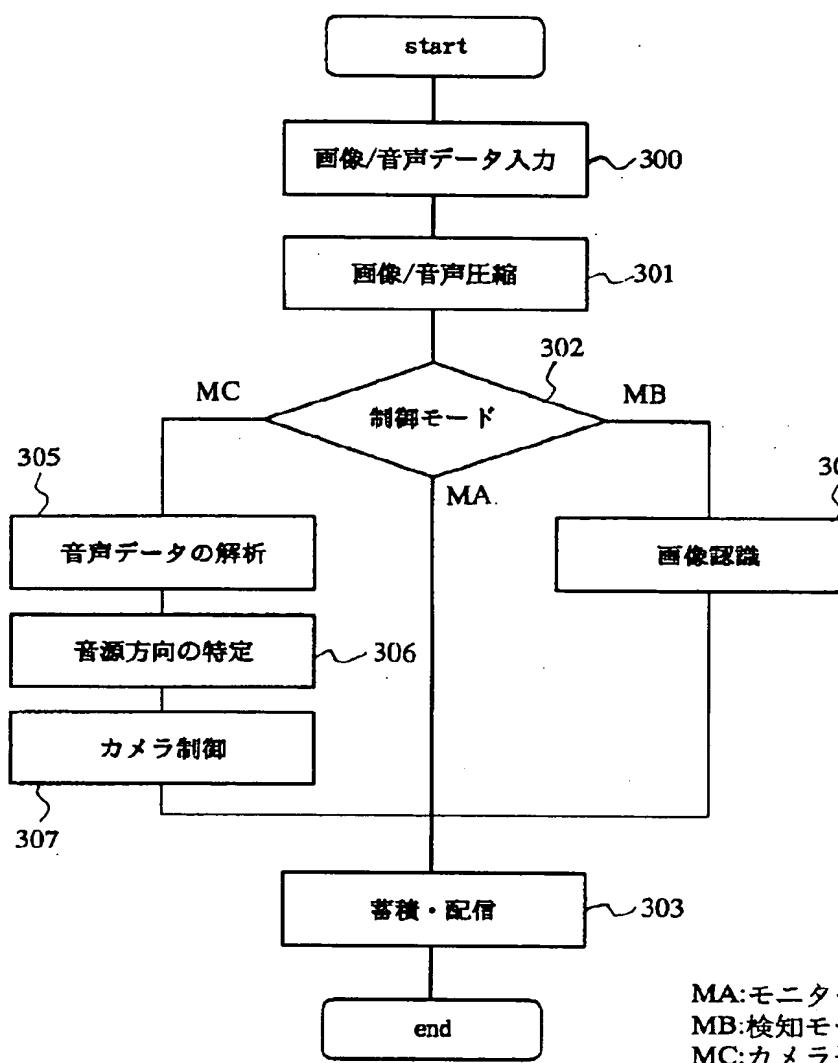
【図1】



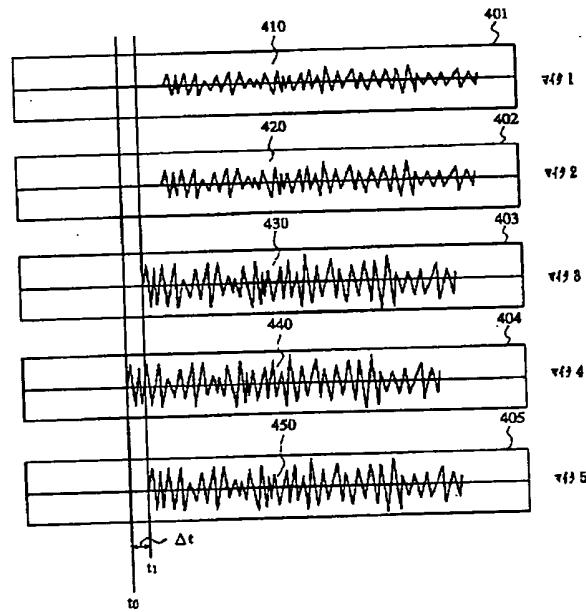
【図2】



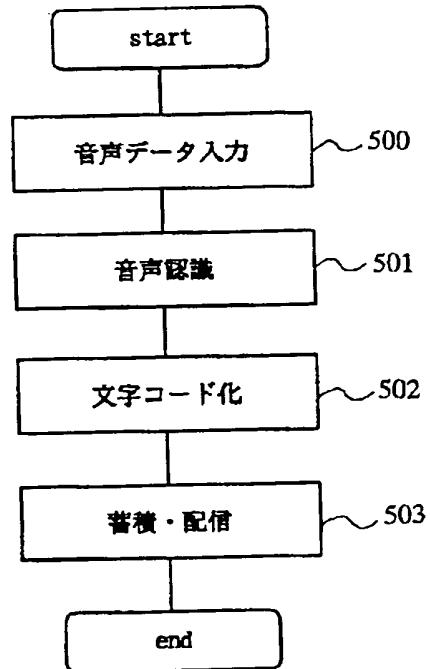
【図3】



【図4】



【図5】



フロントページの続き

(51) Int. Cl. <sup>7</sup>	識別記号	F I	チ-マコ-ト <sup>®</sup> (参考)
G 1 O L	15/28	H 0 4 N	5/225 C
H 0 4 N	5/222		5/232 B
	5/225		C
	5/232	G 1 O L	3/00 5 1 1
			5 5 1 S

F ターム (参考) 5C022 AA01 AB62 AB63 AB65 AC27  
AC71 AC72  
5C054 CA04 CA08 CF01 CF06 CG02  
CG03 CG05 DA06 EA01 EA03  
EA05 FC12 FC13 FF02 HA19  
5C084 AA02 AA07 AA14 AA18 BB31  
CC16 DD02 DD12 EE02 EE05  
FF03 GG17 GG43 GG78  
5C087 AA02 AA03 AA08 BB02 BB46  
BB65 BB74 DD05 DD27 EE06  
EE14 FF01 FF02 FF19 FF30  
GG02 GG59  
5D015 DD02 KK02

**THIS PAGE BLANK (USPTO)**